Copyright © 1975 American Telephone and Telegraph Company
THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 8, October 1975
Printed in U.S.A.

Single-Integration, Adaptive Delta Modulation

By P. CUMMISKEY

(Manuscript received March 11, 1975)

An estimate of optimum performance is derived for a single-integration, adaptive delta modulator. Several simulations of adaptive delta modulators with single integrators have all produced signal-to-noise ratios near or below the estimate.

The derivations presented here indicate that the performance of a single-integration delta modulator is dependent on the correlation between adjacent samples of the input signal and on the probability density function of its derivative. The relationship between the probability density of the derivative of the input signal and optimum performance, in turn, explains why signal-to-noise ratios taken on sine waves are greater than those recorded while processing speech signals.

I. INTRODUCTION

In this paper, an equation is derived for the optimum signal-to-noise ratio (s/n) of a single-integration, adaptive delta modulator. Mean-square quantizing noise is a mathematically tractable quantity which appears to be a reasonably good measure of overall performance. It was felt that an understanding of the relationships between this quantity and the character of the input signal would be useful. The derivations and data presented here all contribute to this end. Other practical considerations, such as subjective evaluation,[1] transmission errors,[2] and tandem encoding,[3] have been discussed elsewhere.

Several simulations[4] of single-integration, adaptive delta modulators on a variety of speech signals have produced s/n's near or below the performance estimate suggested in this paper. It is further suggested that this estimate is very close to the upper bound on the performance of such coders. The s/n formula also provides an explanation of the disparities between s/n's taken on sine waves and those obtained while coding speech signals.

A block diagram of a single-integration, adaptive delta modulator is shown in Fig. 1. At the encoder, the difference between an input

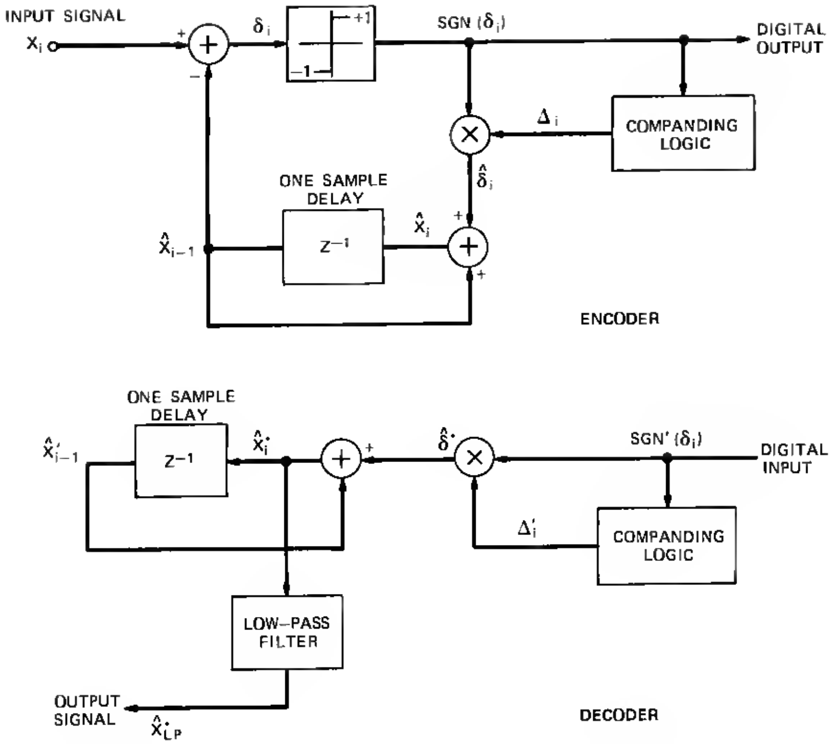

Fig. 1—Single-integration delta modulation.

sample, x_i, and the previous output sample, $\hat{x}_{i-1}$, is quantized to one of two levels and coded. The code symbols, sgn (δ_i) through sgn (δ_{i-N}) (where N may be any positive integer) are then interrogated by the companding logic, and the step size, Δ_{i-1}, is altered before the ith sample is encoded. The quantized approximation to the difference, $\hat{\delta}_i = \Delta_i$ sgn (δ_i), is added to the previous output to obtain the present output sample.

The decoder operates in the same manner as the encoder except that the circuit is driven from the transmission channel rather than from a local comparator. The quantized signal at the decoder, $\hat{x}_i'$, is low-pass filtered to eliminate noise components outside the band of x_i (i.e., frequencies greater than f_{LP}), and a replica of the input signal is thus regenerated at the desampling filter.

The signal-to-noise ratios referred to in this paper were taken in the following manner. First the noise was obtained as shown in Fig. 2 and then the ratio of input signal power to noise power was taken. The technique used by DeJager for sine wave s/n's is described in Ref. 5.

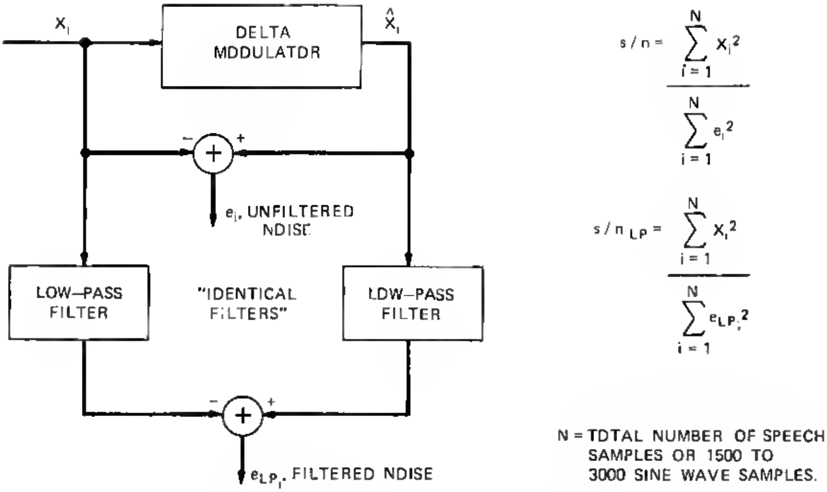

Fig. 2—Quantizing noise measurement.

II. EXACT S/N FORMULAS

The following equations are derived from the diagram in Fig. 1.

$$\delta_i = x_i - \hat{x}_{i-1} \tag{1}$$

$$\hat{x}_i = \hat{x}_{i-1} + \hat{\delta}_i. \tag{2}$$

If the quantizing error is defined as

$$e_i \equiv \hat{x}_i - x_i, \tag{3}$$

then, from (1) and (2), the following relationship holds:

$$e_i = \hat{\delta}_i - \delta_i. \tag{4}$$

From (3), it can be concluded that

$$\hat{x}_i = x_i + e_i, \tag{5}$$

and likewise that

$$\hat{x}_{i-1} = x_{i-1} + e_{i-1}. \tag{6}$$

Therefore, (1) may be rewritten as

$$\delta_i = x_i - x_{i-1} - e_{i-1}. \tag{7}$$

The average power in the prediction error is therefore

$$E(\delta_i^2) = E(x_i^2) + E(x_{i-1}^2) + E(e_{i-1}^2) - 2E(x_i x_{i-1}) - 2E(x_i e_{i-1}) + 2E(x_{i-1}e_{i-1}), \tag{8}$$

where the E functions are expected or average values. It is now noted

that, for quasi-stationary signals,

$$E(x_i^2) = E(x_{i-1}^2),^* \tag{9}$$

and that

$$E(e_{i-1}^2) = E(e_i^2).^* \tag{10}$$

Therefore, eq. (8) may he reduced to

$$\frac{E(\delta_i^2)}{E(x_i^2)} = 2\left[1 - \frac{E(x_i x_{i-1})}{E(x_i^2)} - \frac{E(x_i e_{i-1})}{E(x_i^2)} + \frac{E(x_{i-1} e_{i-1})}{E(x_i^2)}\right] + \frac{E(e_i^2)}{E(x_i^2)}. \tag{11}$$

The s/n at the quantizer is given as

$$\mathrm{s/n}_Q = \frac{E(\delta_i^2)}{E[(\hat{\delta}_i - \delta_i)^2]} = \frac{E(\delta_i^2)}{E(e_i^2)}. \tag{12}$$

The s/n before filtering is defined as

$$\mathrm{s/n} = \frac{E(x_i^2)}{E(e_i^2)}. \tag{13}$$

Note that (11) is equal to (12) divided by (13) or that

$$\frac{E(\delta_i^2)}{E(x_i^2)} = \frac{\mathrm{s/n}_Q}{\mathrm{s/n}}. \tag{14}$$

Hence, by substituting into (11) and transposing terms, an equation for the unfiltered s/n is obtained.

$$\mathrm{s/n} = \frac{\mathrm{s/n}_Q - 1}{2[1 - [E(x_i x_{i-1})]/E(x_i^2) - [E(x_i e_{i-1})]/E(x_i^2) + [E(x_{i-1} e_{i-1})]/E(x_i^2)]}. \tag{15}$$

III. ASSUMPTIONS AND APPROXIMATE FORMULAS

The variance of the prediction error is unknown hecause δ_i contains quantizing noise [see (7)]. Therefore, δ_i cannot be optimally quantized.

No meaningful information can be obtained directly from eqs. (1) through (15) without making some approximations or assumptions about the unknown terms [$\mathrm{s/n}_Q$, $E(x_i e_{i-1})$ and $E(x_{i-1} e_{i-1})$]. Several measurements and simulations taken by the author and others before him support the following assumptions.

(*i*) The optimum step size will yield the same signal-to-noise ratio at the quantizer that can he achieved by quantizing the noise-free part of δ_i (i.e., the derivative of the input signal, $x_i - x_{i-1}$).

* To the extent that (9) and (10) are equations, (15) may be called an equation. Some awkward anomalies exist with regard to eq. (15); however, none of these is relevant to the problem.

(*ii*) The quantizing noise is the same as that generated by optimum quantization of $(x_i - x_{i-1})$, and therefore

$$E[(x_i - x_{i-1})e_{i-1}] = 0. \tag{16}$$

Hence,

$$E(x_i e_{i-1}) - E(x_{i-1}e_{i-1}) = 0. \tag{17}$$

Given the above assumptions, (15) reduces to

$$\text{s/n} = \frac{(\text{s/n}_{qop} - 1)}{2\{1 - [E(x_i x_{i-1})]/E(x_i^2)\}}, \tag{18}$$

where s/n_{qop} is the s/n achieved when $x_i - x_{i-1}$ is optimally quantized.

(*iii*) Finally, in an optimum modulator the quantizing noise spectrum is flat. Then the ratio of overall noise to the inband noise is equal to the ratio of half the sampling frequency to the bandwidth of the input signal.

Hence, the s/n taken on the filtered signal, $\hat{x}'_{LP}$, is equal to the unfiltered s/n multiplied by the ratio of half the sampling frequency to the cutoff frequency of the filter.

$$\text{s/n}_{LP} = \frac{[\text{s/n}_{qop} - 1]\left[\dfrac{f_s}{2f_{LP}}\right]}{2\left[1 - \dfrac{E(x_i x_{i-1})}{E(x_i^2)}\right]}, \tag{19}$$

where f_s is the sampling rate and f_{LP} is the cutoff frequency of the desampling filter or the bandwidth of the input signal.

Equation (19) is identical to Nitadori's signal-to-noise equation[6] for differential PCM. Nitadori cautions against its use in cases where the quantization is coarse, however. In this paper, eq. (19) is derived using somewhat different assumptions which, in fact, do appear to hold for delta modulation.

The validity of the three assumptions given above is the main point of this paper. When these assumptions hold, an important relationship between the amplitude distribution of the derivative of the input signal and s/n performance can be drawn.

IV. RELATIONSHIP BETWEEN S/N_{qop} AND PROBABILITY DENSITY FUNCTION OF $x_i - x_{i-1}$

Paez and Glisson,[7] among others, have shown that the amplitude probability distribution of speech and its derivatives is closely approximated by the gamma distribution. Figure 3 shows that this dis-

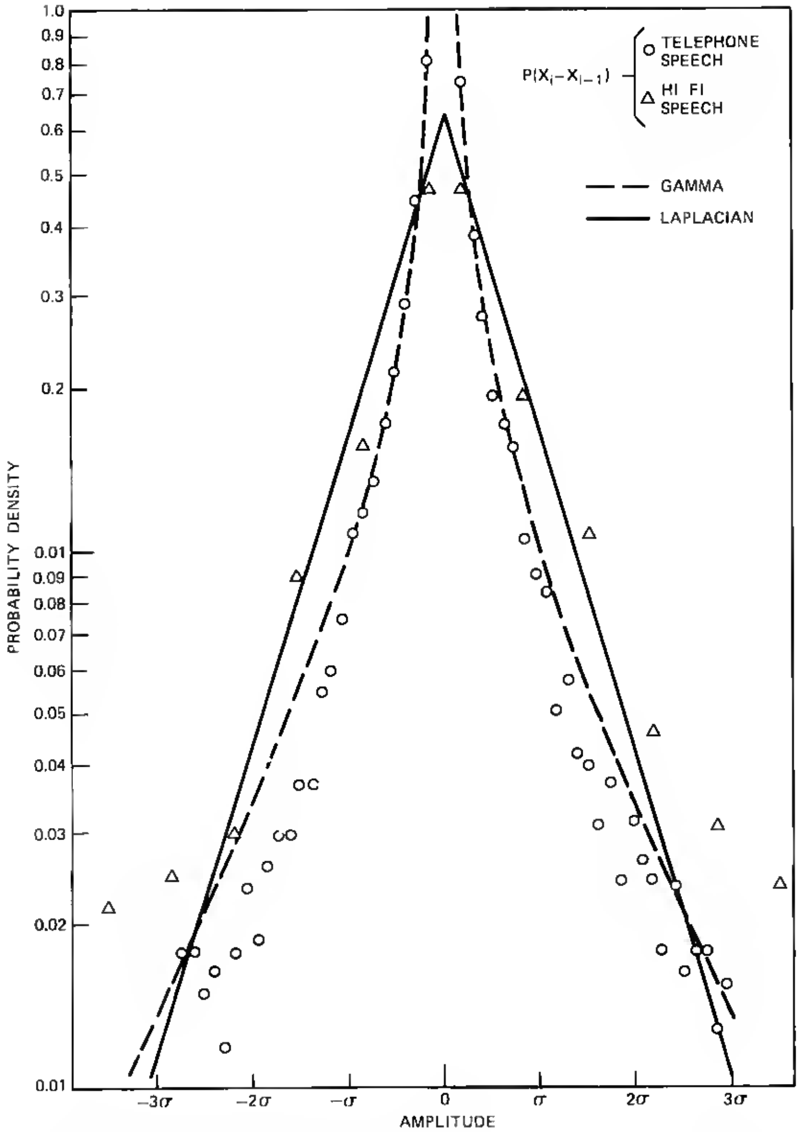

Fig. 3—The amplitude probability density function of $(x_i - x_{i-1})$ as compared with the Laplacian and gamma distributions.

tribution closely approximates the probability distribution of $(x_i - x_{i-1})$ for telephone signals used in my simulations. The distribution of $(x_i - x_{i-1})$, taken on the speech used in Jayant's simulations, lies closer to a Laplacian distribution, however.

Table I — Signal-to-noise ratios of two-level quantizer output

Probability Density Function		s/n$_{qop}$
Gamma	$P(y) = \dfrac{3^{\frac{1}{4}} \exp\left(-\sqrt{3}\|y\|/2\sigma\right)}{\sqrt{8\pi\sigma\|y\|}}$	1.50
Laplacian	$P(y) = \dfrac{1}{\sqrt{2}\sigma} \exp\left(-\dfrac{\sqrt{2}\|y\|}{\sigma}\right)$	2.00
Gaussian	$P(y) = \dfrac{\exp\left(-y^2/2\sigma^2\right)}{\sqrt{2\pi}\sigma}$	2.75
Rectangular	$P(y) = \dfrac{1}{A} \quad -\dfrac{A}{2} < y < \dfrac{A}{2}$ $P(y) = 0 \quad -\dfrac{A}{2} > y > \dfrac{A}{2}$	4.00
Sinusoidal $y = \cos\theta$ or $y = \sin\theta$	$P(y) = \dfrac{1}{\pi\sqrt{1-y^2}} \quad -1 \leqq y \leqq 1$	5.28

Given a distribution that is symmetrical about the origin, the quantization step is optimum when

$$\int_0^\infty (y - \Delta)P(y)dy = 0, \tag{20}$$

where y relates to $(x_i - x_{i-1})$. With the step set at the optimum size, Paez and Glisson, Max,[8] and others have calculated the noise power at the output of a two-level quantizer,

$$E[(\hat{y} - y)^2] = 2\int_0^\infty (y - \Delta)^2 P(y)dy, \tag{21}$$

and achieved the s/n's shown in Table I.

V. COMPARISONS WITH SIMULATIONS

The correlations between adjacent samples was taken on speech obtained using a carbon-button, telephone transducer. Similar data were obtained by N. S. Jayant on speech recorded from a high-fidelity transducer. Both signals were processed by Jayant's adaptive delta modulator with a one-bit memory, where the step size is multiplied by 1.5 if the present and previous code words, sgn (δ_i) and sgn (δ_{i-1}), are alike, or by 0.66 if they differ. In all the simulations, the sampling and desampling filter cutoff frequencies are set at 3.3 kHz, except for the telephone speech recorded at 24 kHz. In this case, the cutoff was

reduced to 3 kHz. The telephone signals sampled at 48 kHz were also encoded hy a single-integration delta modulator designed by D. E. Blahut.[9] Blahut's encoder also performs close to the estimate [eq. (19)], when processing telephone speech. Among the numerous coders tested, no single-integration delta modulator was found that performs significantly hetter than Blahut's or Jayant's.

In Tahle II, performance estimates hased on eq. (19) are compared with the s/n's ohtained using Blahut's and Jayant's delta modulators. To account for difference in prohahility density functions (see Fig. 3), the estimates were made with s/n_{qop} equal to 1.5 for telephone signals, and 2.0 for high-fidelity signals.

The performance estimate given hy eq. (19) is within 3.3 dB of the s/n's obtained in simulations with Jayant's delta modulator. The s/n's taken on Jayant's and Blahut's coders, while processing telephone speech at 48 kHz and 24 kHz, are essentially equal to the estimate. In these cases, the signal level was carefully adjusted until optimum performance was ohtained, then further data were taken to verify eq. (19). (See Tahle III.)

The results shown in Table III lend great support to the approximations made in deriving eq. (19). The noise terms do effectively cancel, leaving a residue that is at least an order of magnitude smaller than the noise-free terms in the denominator of (15) (see lines 5 and 6 in Tahle III). The estimates for noise rejection at the desampling filter and for quantizer performance (s/n_Q) are within 0.8 dB of the figures ohtained in the simulations.

Both coders were simulated with a 60-dB range of step sizes, and hoth were started with the step size equal to the minimum and the

Table II — Performance estimates

s/n_{qop}	Sampling Rate (kHz)	$\frac{E(x_i x_{i-1})}{E(x_i^2)}$	Estimate 10 $\log_{10}$ (s/n_{LP}) (dB)	Delta Modulator Performance (dB)	
				Jayant's	Blahut's
2.0	20	0.989	21.3	18.0	—
*1.5	24	0.957	13.7†	14.5	—
2.0	40	0.997	30.0	28.0	—
*1.5	48	0.9897	22.6	22.9	22.7
2.0	60	0.999	36.5	34.0	—

* Telephone speech: The acoustic-to-electronic response of the new 500-type, stations sets[10] indicates that signal components in the 100-Hz to 3.3-kHz band are differentiated, and that components below 100 Hz are severely attenuated. Hence, correlation between adjacent samples is lower for telephone speech than for high-fidelity speech.

† At 24-kHz sampling, f_{LP} = 3 kHz.

Table III — Verification of eq. (19)

	Sampling Frequency (kHz)	Estimates	Coder Performance	
			Jayant's	Blahut's
s/n_{LP}	24 48	13.7 dB 22.6 dB	14.5 dB 22.9 dB	— 22.7 dB
s/n	24 48	5.92 (i.e., 7.72 dB) 24.3 (i.e., 13.9 dB)	5.95 (i.e., 7.75 dB) 24.3 (i.e., 13.9 dB)	— 26.1 (i.e., 14.2 dB)
Noise rejection at the desampling filter $\approx 10 \log_{10}(f_s/2f_{LP})$	24 48	6.0 dB 8.7 dB	6.8 dB 9.0 dB	— 8.5 dB
s/n_Q	24 48	1.5 1.5	1.531 1.555	— 1.569
$\left[1 - \frac{E(x_i x_{i-1})}{E(x_i^2)}\right]$	24 48	— —	0.0422 0.0103	— 0.0103
$\left[\frac{E(x_i e_{i-1})}{E(x_i^2)} - \frac{E(x_{i-1} e_{i-1})}{E(x_i^2)}\right]$	24 48	0 0	−0.00238 −0.00112	— −0.00060

predictor voltage equal to zero. As the average input signal level was varied over a 40-dB range, it was found that the s/n varied by 3 dB. In either coder, it was found that when performance fell significantly below the estimate (19), the following phenomena were observed:

(*i*) Quantizer performance and unfiltered s/n changed *slightly* (in some cases, these parameters increased in value).
(*ii*) The noise terms no longer effectively canceled.
(*iii*) There was a dramatic reduction in noise rejection at the desampling filter. It appears that when the correlation between the difference signal, $(x_i - x_{i-1})$, and the noise (16) becomes significant, more noise *must* shift into the passband of the desampling filter.

Hence, the approximations used in deriving eq. (19) do appear to describe the optimum condition.

These results have been obtained using both an HP2100A minicomputer and an IBM 370, and therefore are repeatable. Moreover, further validation by others using other encoders is desirable.

VI. SINE WAVE PERFORMANCE

Another interesting check on the theory is the fact that it explains why researchers everywhere achieve much higher s/n with sine wave inputs than with speech signals. DeJager's formula [see eq. (22)] indicates that the s/n taken on a sine wave at any frequency below 3 kHz is greater than the s/n that we predict or obtain for telephone speech.

$$\mathrm{s/n}_{\mathrm{DeJager}} = (0.04)\,\frac{f_s^3}{f^2 \cdot f_{LP}}, \tag{22}$$

where f is the frequency of the input sine wave.

The amplitude probability distribution and $\mathrm{s/n}_{qop}$ for a sine wave were given in Table I. Substitution of the value in Table I into eq. (19) yields an estimate for sine wave s/n's.

$$\mathrm{s/n}_{\text{sine wave}} = \frac{4.28\left(\dfrac{f_s}{2f_{LP}}\right)}{2\left[1 - \dfrac{E(x_i x_{i-1})}{E(x_i^2)}\right]}. \tag{23}$$

Equation (23), in turn, is approximately equivalent to DeJager's formula. This relationship can be shown as follows. Let $x = \sin(2\pi f t)$; then

$$\frac{E(x_i x_{i-1})}{E(x_i^2)} = \frac{\displaystyle\int_0^{1/f} [\sin(2\pi f t)] \cdot \sin(2\pi f t + 2\pi f/f_s)\,dt}{\displaystyle\int_0^{1/f} \sin^2(2\pi f t)\,dt}, \tag{24}$$

or

$$\frac{E(x_i x_{i-1})}{E(x_i^2)} = \cos\left(\frac{2\pi f}{f_s}\right). \tag{25}$$

If the delay angle, $(2\pi f/f_s)$, is sufficiently small, then

$$2\left[1 - \cos\left(\frac{2\pi f}{f_s}\right)\right] \approx 1 - \cos^2\left(\frac{2\pi f}{f_s}\right) \approx \frac{4\pi^2 f^2}{f_s^2}. \tag{26}$$

When (26) is substituted into (23), we obtain something very close to DeJager's formula:

$$\mathrm{s/n}_{\text{sine wave}} = (0.054)\,\frac{f_s^3}{f^2 f_{LP}}. \tag{27}$$

For $f = 800$ Hz, $f_{LP} = 3.3$ kHz, and $f_s = 48$ kHz, estimates of 33.5 and 34.7 dB are obtained using (22) and (27). Under these same conditions, signal-to-noise readings of 26 to 27 dB were obtained in simula-

tions of Jayant's delta modulator. Under similar conditions, DeJager[5] obtained a maximum s/n of about 30 dB on a linear delta modulator.

VII. CONCLUSIONS

The optimum performance of Blahut's and Jayant's delta modulators is very close to the estimate, (19), when processing speech signals. Further experimentation with step-size compandors, without a change in the prediction technique, will not produce significantly higher signal-to-noise ratios. Equation (19) applies to a delta modulator with a single, ideal integrator; therefore, it does not preclude improvements through the use of fixed, higher-order networks.

In addition, it has been shown that delta modulator performance is dependent on the amplitude probability distribution of the derivative of the input signal. This dependence should be tested on a variety of signals and probability density functions. The theory also implies that a relationship exists between the amplitude distributions of differential waves at the input and optimum s/n, when higher-order networks are used.

Finally, I wish to call attention to the fact that the s/n performance of a delta modulator is significantly less for telephone signals than for low-pass filtered, high-fidelity signals, or for sine waves.

VIII. ACKNOWLEDGMENTS

I would like to thank D. E. Blahut and J. A. Miller for supplying me with working programs. Acknowledgments must also be extended to D. J. Hunsberger for instructing me on the use of the Hewlett Packard 2100A computer. H. J. Fletcher, D. E. Blahut, and N. S. Jayant have engaged in many useful discussions from which I have drawn insights. N. S. Jayant must also be thanked for supplying data related to his simulations. W. R. Daumer's efforts in collecting data for subjective tests have also provided useful information. I should also like to thank N. S. Jayant and C. E. Nahabedian for their help with the manuscript.

REFERENCES

1. N. S. Jayant and A. E. Rosenburg, "The Preference of Slope-Overload to Granularity in Delta Modulation of Speech," B.S.T.J., *50*, No. 10 (December 1971), pp. 3117–3125.
2. P. T. Nielsen and J. L. Flanagan, "Effects of Channel Errors and Adaptive Delta Modulation Systems," unpublished work.
3. N. S. Jayant and K. Shipley, "Multiple Delta Modulation of a Speech Signal," Proc. IEEE (lett.), *59* (September 1971), p. 1382.
4. N. S. Jayant, "An Adaptive Delta Modulator with a One-Bit Memory," B.S.T.J., *49*, No. 3 (March 1970), pp. 321–342.

5. F. DeJager, "Delta Modulation, A Method of PCM Transmission Using a One-Unit Code," Phillips Research Report, *7* (1952), pp. 442–446.
6. K. Nitadori, "Statistical Analysis of ΔPCM," Electronics and Communications in Japan, *48*, No. 2 (February 1965), pp. 17–26.
7. M. D. Paez and T. H. Glisson, "Minimum Mean Square Quantization in Speech, PCM and DPCM Systems," IEEE Trans. on Comm., *20*, No. 2 (April 1972), pp. 225–230.
8. J. Max, "Quantization for Minimum Distortion," IRE Trans. Information Theory, *IT-6* (March 1960), pp. 7–12.
9. D. E. Blahut, U.S. Patent 3,784,922.
10. A. H. Iglis and W. L. Tuffnell, "An Improved Telephone Set," B.S.T.J., *30*, No. 1 (January 1951), pp. 239–270.